НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Aggregation of journal rankings: an example of application of social choice in scientometrics

Andrey Subochev, Fuad Aleskerov, Vladimir Pislyakov

Motivation

How to construct a reasonably good representation of the set of rankings which are based on bibliometric indicators?

Selected indicators

Indicator	*Database*	*Year*	*Publication window, years*	*Weighted*	*Size-dependent*
Impact factor	WoS/JCR	2011	2	No	No
5-year impact factor	WoS/JCR	2011	5	No	No
Immediacy index	WoS/JCR	2011	1	No	No
Article influence	WoS/JCR	2011	5	Yes	No
Hirsch index	WoS	2007–2011 (papers and citations)	5	No	Yes
SNIP	Scopus	2011	3	No	No
SJR	Scopus	2011	3	Yes	No

- Economics: 212 journals
- Management: 93
- Political Science: 99

Rank correlations

Share of inversions, % (economic journals)

	Impact factor	5-year impact factor	Immediacy index	Article influence	Hirsch index	SNIP	SJR
Impact factor		**8.46**	24.59	18.13	15.45	15.09	14.23
5-year impact factor	**8.46**		24.25	13.72	13.15	13.66	12.20
Immediacy index	24.59	24.25		26.00	25.57	**27.01**	25.25
Article influence	18.13	13.72	26.00		17.15	16.31	15.50
Hirsch index	15.45	13.15	25.57	17.15		18.47	15.05
SNIP	15.09	13.66	**27.01**	16.31	18.47		17.28
SJR	14.23	12.20	25.25	15.50	15.05	17.28	

Social choice

X – the *general set* of alternatives

A – the *feasible set* of alternatives: $A \subseteq X \wedge A \neq \varnothing$. The feasible set is a variable.

N – the *society* (a group of voters or a panel of experts)

$u_i(x)$ – the *utility* of alternative $x \in X$ for voter $i \in N$, $u_i(x): X \to \mathbb{R}$

$$u_i(y) > u_i(x) \Leftrightarrow \text{voter } i \text{ strictly prefers } y \text{ to } x$$

$U = \{ u_i(x) \mid i \in N \}$ – the profile of utility functions

R – (*weak*) *social preferences*, $R \subseteq X \times X$

R is presumed to be complete: $\forall\, x \in X, \forall\, y \in X, (x, y) \in R \vee (y, x) \in R$

P – *strict social preferences*, $P \subseteq R$: $(x, y) \in P \Leftrightarrow ((x, y) \in R \wedge (y, x) \notin R)$

It is presumed that $R = R(P)$ and $P = P(U)$.

Axioms of aggregation

Aggregation rule $R=R(U)$

- ***Completeness***: all alternatives are comparable, $xR(U)y \vee yR(U)x$
- ***Transitivity***: $(xR(U)y \wedge yR(U)z) \Rightarrow xR(U)z$
- ***Neutrality***: the rule treats all alternatives equally
- ***Anonymity***: the rule treats all aggregated rankings equally
- ***Strong Pareto principle***: if x Pareto-dominates y, then xPy
- ***Full domain***: the rule can be applied in all cases, i.e. to any utility profile U
- ***Independence of irrelevant utilities***: $\forall A \subseteq X$, $P(U)|_A=P(U|_A)$
- ***Ordinality***: if utility profiles U and U' are such that $\forall x, y \in A \subseteq X$, $\forall i \in N$, $u_i(x) > u_i(y) \Leftrightarrow u'_i(x) > u'_i(y)$, then $R(U|_A)=R(U'|_A)$ for any such $A \subseteq X$.

The majority rule and the majority relation P (formal definitions and representations)

N – the set of indicators; $u_k(x)$ – the value of indicator k for journal x

The majority rule

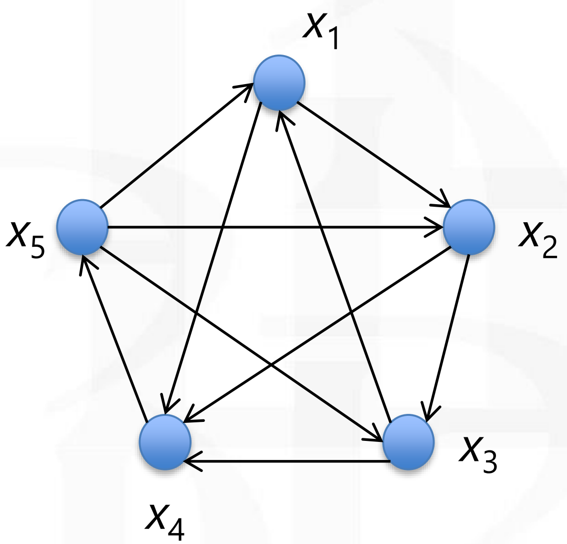
x is better than $y \Leftrightarrow \#\{k \in N \mid u_k(x) > u_k(y)\} > \#\{k \in N \mid u_k(y) > u_k(x)\}$

P – **the majority relation**: $(x, y) \in P \Leftrightarrow x$ is majority-preferred to y

$\mathbf{M}=[m_{ij}]$ – matrix representing P: $m_{xy}=1 \Leftrightarrow (x, y) \in P$, $m_{xy}=0 \Leftrightarrow (x, y) \notin P$

	x_1	x_2	x_3	x_4	x_5
x_1	0	1	0	1	0
x_2	0	0	1	1	0
x_3	1	0	0	1	0
x_4	0	0	0	0	1
x_5	1	1	1	0	0

Tournament matrix **M**



Majority digraph

Why the Majority rule? An axiomatic argument

- ***Completeness***
- ~~***Transitivity***~~
- ***Neutrality***
- ***Anonymity***
- ***Strong Pareto principle***
- ***Full domain***
- ***Independence of irrelevant utilities***
- ***Ordinality***
- ***Strict Cardinal Monotonicity***
- ***Positive responsiveness***
- ***Computational Simplicity***

Why the majority rule? An axiomatic argument

- ***Strict Cardinal Monotonicity***: if utility profiles U and U' are such that $\forall i \in N, u'_i(x) \geq u_i(x) \wedge u'_i(y) = u_i(y)$, then $xP(U)y \Rightarrow xP(U')y$ and $xR(U)y \Rightarrow xR(U')y$

- ***Positive responsiveness***: if utility profiles U and U' are such that $\exists j \in N$: $(u_j(x) < u_i(y) \wedge u'_j(x) \geq u'_j(y)) \vee (u_j(x) = u_j(y) \wedge u'_j(x) > u'_j(y))$ and $\forall i \in N \setminus \{j\}, u'_i(x) = u_i(x) \wedge u'_i(y) = u_i(y)$ and $xR(U)y$ and $yR(U)x$ then $xP(U')y$

- ***Computational simplicity***: there exists a polynomial algorithm for computing $R(U)$.

The majority rule (example)

№	Journal	IF	5-IF	Immediacy index	Article influence	Hirsch	SNIP	SJR
1	Explorations in Economic History	**0.935**	0.898	**0.541**	0.772	7	**1.768**	**0.036**
2	Review of Income and Wealth	0.805	**1.103**	0.205	**0.850**	**9**	1.712	0.034

$$4 > 3$$

J_1 is better than J_2

The Condorcet paradox

Journal	IF	5-IF	Immediacy index	Article influence	Hirsch	SNIP	SJR
Explorations in Economic History	**0.935**	0.898	**0.541**	0.772	7	**1.768**	**0.036**
Review of Income and Wealth	**0.805**	**1.103**	**0.205**	0.850	**9**	**1.712**	0.034
Scandinavian Journal of Economics	0.514	**1.070**	0.150	**1.310**	**8**	1.426	**0.043**

J_1 is better than J_2 (4 > 3)

J_2 is better than J_3 (5 > 2)

J_3 is better than J_1 (4 > 3)

Numbers of 3-, 4- and 5-step *P*-cycles and ties

	3-step cycles	***4-step cycles***	***5-step cycles***	***Tied pairs***	***All pairs***
Economics	2446	22427	226103	197	22366
Management	203	787	3254	33	4278
Political Science	149	430	1344	73	4851

Majority-rule-based ranking procedures

- The ***Copeland rule***
  (ranking by the number of victories won in a tournament *P*)
  Version ***2*** (a tie is counted as a loss)
  Version ***3*** (a tie is counted as a victory)

- A sorting based on a ***tournament solution***,
  which determines the winners of a tournament *P*
  The best alternatives (the "winners") are determined by
  the ***uncovered set*** *UC*
  the ***minimal externally stable*** set *MES*

- Ranking the nodes of a digraph representing *P* by ***Markovian*** random walk method

The Copeland rule. Axiomatic analysis

- ***Completeness***
- ***Transitivity***
- ***Neutrality***
- ***Anonymity***
- ***Strong Pareto principle***
- ***Full domain***
- ~~***Independence of irrelevant utilities***~~
- ***Ordinality***
- ***Strict Cardinal Monotonicity***
- ~~***Positive responsiveness***~~
- ***Computational Simplicity***

- ***Weak Arrowian Independence irrelevant alternatives***

The Copeland rule. Axiomatic analysis (continued)

- ***Arrowian Independence of irrelevant alternatives***

AIIA ⇔ Independence of irrelevant utilities ∧ Ordinality

$$\forall A \subseteq X,\ \forall x, y \in A,\ \forall i \in N,\ xR_iy \Leftrightarrow xR'_iy \ \wedge\ xP_iy \Leftrightarrow xP'_iy$$

$$\Rightarrow$$

$$xR(U|_A)y \Leftrightarrow xR(U'|_A)y \ \wedge\ xP(U|_A)y \Leftrightarrow xP(U'|_A)y.$$

- ***Weak Arrowian Independence of irrelevant alternatives***

Suppose the feasible set A is fixed. Then $\forall x, y \in A$,

$$\forall i \in N,\ \forall z \in A,\ xR_iz \Leftrightarrow xR'_iz \wedge xP_iz \Leftrightarrow xP'_iz \wedge yR_iz \Leftrightarrow yR'_iz \wedge yP_iz \Leftrightarrow yP'_iz$$

$$\Rightarrow$$

$$xR(U|_A)y \Leftrightarrow xR(U'|_A)y \ \wedge\ xP(U|_A)y \Leftrightarrow xP(U'|_A)y.$$

The sorting by *MES*. Axiomatic analysis

- ***Completeness***
- ***Transitivity***
- ***Neutrality***
- ***Anonymity***
- ~~***Strong Pareto principle***~~
- ***Full domain***
- ~~***Independence of irrelevant utilities***~~
- ***Ordinality***
- ~~***Strict Cardinal Monotonicity***~~
- ~~***Positive responsiveness***~~
- ***Computational Simplicity***

- ***Weak Pareto principle***

  if x Pareto-dominates y, then xRy
- ***Independence of classes of irrelevant alternatives***
- ***Cardinal Monotonicity***: if profiles

U, U' are s.t. $\forall i \in N$, $u'_i(x) \geq u_i(x) \wedge u'_i(y) = u_i(y)$, then $xR(U)y \Rightarrow xR(U')y$

The sorting by *MES*. Axiomatic analysis (continued)

- ***Idempotency***: $\forall A$, $S(S(A))=S(A)$.
- ***The Aizerman-Aleskerov condition***: $\forall A, \forall B, S(A)\subseteq B\subseteq A \Rightarrow S(B)\subseteq S(A)$.
- ***Nash Independence of irrelevant alternatives*** (*I. of outcasts*):
  $\forall A, \forall B, S(A)\subseteq B\subseteq A \Rightarrow S(B)=S(A)$.

NIIA $\Leftrightarrow$ ***Idempotency*** $\wedge$ ***the Aizerman-Aleskerov condition***

If a ranking rule *R* is a sorting based on a tournament solution *S* then

R satisfies *Independence of classes of irrelevant alternatives* and

(Cardinal/Ordinal) *Monotonicity* if *S* satisfies the *Nash IIA.*

MES satisfies the *Nash IIA.*

The sorting by *UC*. Axiomatic analysis

- ***Completeness***
- ***Transitivity***
- ***Neutrality***
- ***Anonymity***
- ***Strong Pareto principle***
- ***Full domain***
- ~~***Independence of irrelevant utilities***~~
- ***Ordinality***
- ~~***Monotonicity***~~
- ~~***Positive responsiveness***~~
- ***Computational Simplicity***

Rank correlations (continued)

Kendall τ_b (economic journals)

	Impact factor	5-year impact factor	Immediacy index	Article influence	Hirsch index	SNIP	SJR	Copeland (2)	Copeland (3)	*UC*	*MES*	Marcovian
Impact factor	1.000	0.830	0.503	0.637	0.654	0.698	0.700	0.834	0.831	0.834	0.835	0.819
5-year IF	0.830	1.000	0.510	0.725	0.702	0.726	0.741	0.903	0.904	0.906	0.896	0.891
Immediacy index	0.503	0.510	1.000	0.475	0.442	0.454	0.472	0.550	0.551	0.556	0.578	0.560
Article influence	0.637	0.725	0.475	1.000	0.620	0.673	0.674	0.766	0.769	0.777	0.785	0.769
Hirsch index	0.654	0.702	0.442	0.620	1.000	0.592	0.650	0.738	0.737	0.737	0.747	0.729
SNIP	0.698	0.726	0.454	0.673	0.592	1.000	0.638	0.759	0.759	0.767	0.775	0.750
SJR	0.700	0.741	0.472	0.674	0.650	0.638	1.000	0.792	0.790	0.800	0.797	0.775
Copeland (2)	0.834	0.903	0.550	0.766	0.738	0.759	0.792	1.000	0.990	0.970	0.950	0.956
Copeland (3)	0.831	0.904	0.551	0.769	0.737	0.759	0.790	0.990	1.000	0.969	0.950	0.959

Formal analysis of correlations

Kendall τ_b (economic journals)

	IF	5-IF	Immediacy	AI	Hirsch	SNIP	SJR
5-year IF	**0.830**	**1.000**	0.510	0.725	0.702	0.726	0.741
Markovian	0.819	0.891	**0.560**	**0.769**	**0.729**	**0.750**	**0.775**

The Markovian ranking majority-dominates the ranking based on 5-IF

ERGO

The Markovian ranking represents the set of seven single-indicator-based rankings better than the ranking based of 5-year impact factor

Voting matrix (economic journals)

	Impact factor	5-year impact factor	Immediacy index	Article influence	Hirsch index	SNIP	SJR	Copeland (2)	Copeland (3)	*UC*	*MES*	Marcovian
Impact factor		1	6	6	6	5	5	1	1	1	1	1
5-year IF	6		6	6	6	6	6	1	1	1	1	2
Immediacy index	1	1		1	1	1	1	1	1	1	1	1
Article influence	1	1	6		5	4	3	1	1	1	1	1
Hirsch index	1	1	6	2		2	1	1	1	1	1	1
SNIP	2	1	6	3	5		1	1	1	1	1	1
SJR	2	1	6	4	6	6		1	1	1	1	1
Copeland (2)	6	6	6	6	6	6	6		3	1	1	5
Copeland (3)	6	6	6	6	6	6	6	4		0	1	5
UC	6	6	6	6	6	6	6	6	7		2	6
MES	6	6	6	6	6	6	6	6	6	5		7
Markovian	6	5	6	6	6	6	6	2	2	1	0	

Tournament matrix and the Copeland scores (economic journals)

	Impact factor	5-year impact factor	Immediacy index	Article influence	Hirsch index	SNIP	SJR	Copeland (2)	Copeland (3)	*UC*	*MES*	Marcovian	**Copeland score**
Impact factor		0	1	1	1	1	1	0	0	0	0	0	**5**
5-year IF	1		1	1	1	1	1	0	0	0	0	0	**6**
Immediacy index	0	0		0	0	0	0	0	0	0	0	0	**0**
Article influence	0	0	1		1	1	0	0	0	0	0	0	**3**
Hirsch index	0	0	1	0		0	0	0	0	0	0	0	**1**
SNIP	0	0	1	0	1		0	0	0	0	0	0	**2**
SJR	0	0	1	1	1	1		0	0	0	0	0	**4**
Copeland (2)	1	1	1	1	1	1	1		0	0	0	1	**8**
Copeland (3)	1	1	1	1	1	1	1	1		0	0	1	**9**
UC	1	1	1	1	1	1	1	1	1		0	1	**10**
MES	1	1	1	1	1	1	1	1	1	1		1	**11**
Markovian	1	1	1	1	1	1	1	0	0	0	0		**7**

The rankings of rankings

rank	Economics	Man. Sc.	Pol. Sc.	Previous results (2008)
1	*MES*	*MES*	*MES*	*UC*
2	*UC*	*UC*	*UC*	*MES*
3	Copeland 3	Copeland 2	Copeland 3	Copeland 3
4	Copeland 2	Copeland 3	Copeland 2	Copeland 2
5	Markovian	Markovian	Markovian	Markovian
6	5-IF	5-IF	5-IF	IF
7	IF	SNIP	Hirsch	5-IF
8	SJR	Hirsch		SJR
9	AI	AI	AI / IF / SJR	
10	SNIP	SJR		AI / Hirsch / SNIP
11	Hirsch	IF	SNIP	
12	Immediacy	Immediacy	Immediacy	Immediacy

The share of coinciding pairs r

What if we change the measure of correlation?

Let us replace τ_b by the share of coinciding pairs r (a percentage of pairs ranked in the same way in both rankings).

r = 50% means two rankings do not correlate.

The rankings of rankings (Economics)

rank	Compared by	
	τ_b	r
1	*MES*	Copeland 3
2	*UC*	Copeland 2
3	Copeland 3	Markovian
4	Copeland 2	*UC*
5	Markovian	5-IF
6	5-IF	IF
7	IF	*MES*
8	SJR	AI
9	AI	SNIP
10	SNIP	SJR
11	Hirsch	Hirsch
12	Immediacy	Immediacy

The rankings of rankings (Management Science)

rank	Compared by	
	τ_b	*r*
1	*MES*	Copeland 3
2	*UC*	Copeland 2
3	Copeland 2	Markovian
4	Copeland 3	*UC*
5	Markovian	5-IF
6	5-IF	*MES*
7	SNIP	SNIP
8	Hirsch	AI
9	AI	IF / Hirsch / SJR
10	SJR	
11	IF	
12	Immediacy	Immediacy

The rankings of rankings (Political Science)

rank	Compared by	
	τ_b	*r*
1	*MES*	Copeland 3 / Copeland 2 / Markovian
2	*UC*	
3	Copeland 3	
4	Copeland 2	*UC*
5	Markovian	5-IF
6	5-IF	*MES*
7	Hirsch	AI
8	AI / IF / SJR	IF
9		SNIP
10		SJR
11	SNIP	Hirsch
12	Immediacy	Immediacy

The number of ranks

	Number of journals	Impact factor	5-year impact factor	Immediacy index	Article influence	Hirsch index	SNIP	SJR	Copeland (2)	Copeland (3)	*UC*	*MES*	Marcovian
Economics	**212**	200	207	159	204	30	201	65	135	139	59	37	211
Management	**93**	90	92	84	91	30	92	41	68	69	42	33	93
Political Sc.	**99**	95	98	72	95	19	97	28	69	66	42	36	97

Conclusions

- The rankings based on popular bibliometric indicator strongly and positively correlate with each other, but there always is a non-negligible percentage of contradictions.
- To construct a good representation of the set of single-indicator-based rankings one may use a majority-rule-based rank aggregation procedure.

Publications

1. Subochev, A., Aleskerov, F., Pislyakov, V. 2018. Ranking journals using social choice theory methods: A novel approach in bibliometrics. *Journal of Informetrics*, 12(2), 416–429.
2. Subochev A., Pislyakov V. 2018. With or without h-index? Comparing aggregates of rankings based on seven popular bibliometric indicators. *Proceedings of the 23D International Conference on Science and Technology Indicators.* Leiden: Universiteit Leiden-CWTS. P. 1135-1143. ISBN: 978-90-9031204-0.
3. Aleskerov, F., Pislyakov, V., Subochev, A. 2014. Ranking Journals in Economics, Management and Political Science by Social Choice Theory Methods. WP BRP 27/STI/2014. Moscow: HSE.

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Thank you

for

your attention!

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Спасибо за внимание!

ul. Myasnitskaya, 20
Moscow, Russia, 101000
Phone: (495) 621-7983, Fax: (495) 628-7931
www.hse.ru